

Ю. А. Бабкин, А. М. Елизаров

Казанский государственный университет

культуры и искусств,

Казанский (Приволжский) федеральный университет,

yurionbabkin@yandex.ru, elizarov@ksu.ru

ИНФОРМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССА ОБРАЩЕНИЯ К ЭЛЕКТРОННЫМ ИНФОРМАЦИОННЫМ РЕСУРСАМ НАУЧНОЙ БИБЛИОТЕКИ ИМ. Н.И. ЛОБАЧЕВСКОГО

В настоящее время всё большее распространение в мире получает статистический анализ научной и образовательной деятельности (см., например, [1 – 5]). Одно из наиболее важных и хорошо развитых направлений такого анализа составляют выявление и изучение статистических закономерностей производства, поиска и использования информации. Это направление выделилось в отдельную отрасль, названную информетрией [6, 7]. Информетрический анализ играет в настоящее время все большую роль в развитии электронных библиотек в различных областях знания и оценке публикационной активности в сфере образования и научной деятельности.

Информетрическое моделирование информационных процессов заключается в выявлении эмпирических закономерностей, наблюдаемых в значительном числе информационных процессов, выражении их в строгой математической форме и распространении построенных моделей на другие процессы, однотипные исследованным. Такое моделирование, как правило, базируется на известных информетрических законах. Поэтому актуальной проблемой информетрического моделирования, проводимого на конкретном информационном процессе, явля-

ется определение, насколько применим к рассматриваемому процессу тот или иной информетрический закон, насколько корректно и точно он выполняется и каковы особенности его действия в данном случае. Отметим, что Л. Эгге [8] (см. также [7]) выделил класс так называемых "информационных процессов производства" (Information Production Process, IPP) и признаки принадлежности к нему конкретных информационных процессов. Любые из них при информетрическом моделировании априори можно считать сходными, и возникает задача выбора подходящего информетрического закона, на базе которого будет строиться соответствующая математическая модель. Другой задачей является определение параметров нестационарной математической модели с целью наилучшего приближения анализируемых экспериментальных данных.

Одним из IPP является процесс обращения читателей к электронным изданиям. При этом современные средства анализа процесса обращения к документальным онлайн-базам данных позволяют с высокой степенью достоверности получить статистические показатели чтения пользователями электронных изданий, чтобы в дальнейшем применять их в процессе моделирования. Кроме того, современные библиотеки все активнее используют электронные информационные ресурсы и научные БД для удовлетворения читательского спроса, что, в частности, обусловлено необходимостью дополнить традиционные методы обслуживания читателей оперативной и актуальной информацией, точнее определить информационные потребности последних, а также оценить экономическую эффективность использования электронных информационных ресурсов.

Методология применения основных информетрических за-

конов (Брэдфорда, Парето, Леймкулера, Лотки, Ципфа) для математического моделирования процесса обращения читателей к электронным изданиям разработана В.В. Писляковым [9]. Отличительной чертой этого исследования является попытка переноса целого комплекса информетрических моделей из традиционной "печатной" среды в среду электронную. Эта попытка базировалась на данных по использованию электронных ресурсов в Национальном исследовательском университете "Высшая школа экономики" за 2004 год, а результаты ее оказались вполне успешными. Настоящая работа является продолжением исследований [9] и нацелена на проверку применимости закона Брэдфорда к моделированию процесса потребления научной информации. В качестве экспериментальной площадки использовались научные базы данных издательства Elsevier, содержащиеся в Научной библиотеке им. Н.И. Лобачевского Казанского (Приволжского) федерального университета и охватывающие период 2004 – 2011 годов. Генеральная совокупность составляла 2493 наименований журналов. Выборочная совокупность была сформирована на основании критериев, указанных в [9]: по числу обращений к полным текстам статей в журналах. В исследуемой выборочной совокупности присутствовали только источники (журналы), к которым были обращения (1723 из 2493) с открытием 587502 полнотекстовых статей из них. Журналы идентифицировались по ISSN, не востребованные ни разу журналы полностью исключались из рассмотрения.

Проведено моделирование процесса обращения к электронным информационным ресурсам при помощи закона Брэдфорда в двух формах – "вербальной" и графической (их описание см. в [9]). На этой основе построена модель обращения к элек-

тронным источникам – реальные данные разбивались на "зоны Брэдфорда" согласно математической формулировке закона, различными методами был проведен подбор эмпирических коэффициентов, а также построена кривая Брэдфорда ("библиограф"). Установлено, что закон Брэдфорда не выполняется в своей классической "вербальной" формулировке, а график обращения к источникам не соответствует классическому виду кривой Брэдфорда. Предприняты попытки выделения "информационного ядра" на основе имеющегося экспериментального массива с применением модели Брэдфорда. Ядро выделялось двумя методами, использованными в [9], – кубической аппроксимацией кривой Брэдфорда с нахождением точки перегиба и с привлечением математической формулировки закона для определения границы квазипрямолинейного участка библиографа. При проведении этих вычислений установлено, что зависимость кумулятивного числа открытых статей в S журналах от $\ln S$ лучше всего аппроксимируется, как и в [9], полиномом 3-й степени. Достоверность полученных данных подтвердил коэффициент детерминации (R -квадрат), равный 0,9953. Кроме того, в отличие от [9] методику подбора параметров модели Брэдфорда, предложенную Л. Эгге, использовать в полной мере не удалось, поскольку увеличение числа групп Брэдфорда повлекло за собой уменьшение первой зоны, а при числе зон, равном 4, размер первой зоны оказался меньше 1, что невозможно по определению.

Таким образом, результаты, полученные на основе экспериментальных данных об обращении к электронным информационным ресурсам в Научной библиотеке им. Н.И. Лобачевского, подтвердили, как и в [9], применимость закона Брэдфорда при информетрическом моделировании.

Л И Т Е Р А Т У Р А

1. Гохберг Л.М. *Статистика науки*. – М.: ТЕИС, 2003. – 478 с.
2. *Main science and technology indicators*. – OECD, 2008. – V. 2008/1. – 105 p.
3. *National science board. Science and engineering indicators*. – Arlington, VA: National Science Foundation, 2008. – V. 1 – 2.
4. *Индикаторы науки: 2010*. Статистический сборник. Под ред. Л.М. Гохберга, Я.И. Кузьминова, К.Э. Лайкама и др. – М.: НИУ Высшая школа экономики, 2011. – 367 с.
5. *Индикаторы образования: 2011*. Статистический сборник. Под ред. Л.М. Гохберга, Н.В. Ковалёва, Я.И. Кузьминова и др. – М.: НИУ Высшая школа экономики, 2011. – 264 с.
6. Горькова В.И. *Информетрия (количественные методы в научно-технической информации)* // Итоги науки и техники. Сер. Информатика. – М.: ВИНТИ, 1988. – Т. 10. – 328 с.
7. Egghe L., Rousseau R. *Introduction to informetrics: quantitative methods in library, documentation and information science*. – Amsterdam e. a.: Elsevier Science Publishers, 1990. – P. 292-313.
8. Egghe L. *The duality of informetric systems with applications to the empirical laws* // J. of Information Science. – 1990. – V. 16, No 1. – P. 17-27.
9. Писляков В.В. *Моделирование процесса обращения к электронным информационным источникам на основе информетрического закона Брэдфорда* // Уч. записки Казанского государственного университета. Серия Физико-математические науки. – 2007. – Т. 119, Кн. 2. – С. 116-127.